



**eResearch: the open access repository of the  
research output of Queen Margaret University, Edinburgh**

This is an author-formatted version of a document published as:

Stuart-Smith, Jane and Lawson, Eleanor and Scobbie, James M (2014)  
*Derhoticisation in Scottish English: a sociophonetic journey*. In: *Advances in Sociophonetics*. John Benjamins, Amsterdam, pp. 59-96. ISBN  
9789027234957

**Accessed from:**

<http://eresearch.qmu.ac.uk/3084/>

**Repository Use Policy**

The full-text may be used and/or reproduced, and given to third parties for personal research or study, educational or not-for-profit purposes providing that:

- The full-text is not changed in any way
- A full bibliographic reference is made
- A hyperlink is given to the original metadata page in eResearch

eResearch policies on access and re-use can be viewed on our Policies page:  
<http://eresearch.qmu.ac.uk/policies.html>

**Copyright © and Moral Rights for this article are retained  
by the individual authors and/or other copyright owners.**

<http://eresearch.qmu.ac.uk>

## **Derhoticisation in Scottish English: A sociophonetic journey**

Jane Stuart-Smith\*, Eleanor Lawson<sup>§</sup>, and James M. Scobbie<sup>§</sup>

\* *English Language/Glasgow University Laboratory of Phonetics, University of Glasgow*

<sup>§</sup> *CASL Research Centre, Queen Margaret University Edinburgh*

### ***1. Introduction***<sup>1</sup>

This paper presents the concrete example of the rewards of a sociophonetic journey by focusing on an area which is particularly rich and informative – fine-grained variation in Scottish English coda /r/. We synthesize the results of some 15 years of research, including our current work in progress, with those of previous studies, and provide a sociophonological account of variation and change in this feature. This forces us to consider carefully the complex relationships between auditory, acoustic, and articulatory descriptions of (socially structured) speech. Our research also raises questions about speakers’ mental representations of such information.

We begin by summarizing observations on coda /r/ in Scottish English across the twentieth century, which reveal a socially-constrained, long-term process of derhoticisation. Then we consider the most recent evidence for derhoticisation from different perspectives in order to learn more about the nature and mechanism of the change. We look at the linguistic and social factors involved (sections 2 and 3); the views from the listener (section 4); the acoustics of derhoticisation (section 5); and insights from a socio-articulatory corpus collected and analysed using Ultrasound Tongue Imaging (section 6). Finally we discuss the implications of our results for representation, by analysts, and for speaker-hearers in this community.

---

<sup>1</sup> Jane Stuart-Smith is grateful to audiences at the Workshop on Sociophonetics (Pisa), and at seminars at the Universities of Oxford, Stanford, and Berkeley, for their feedback on earlier versions of this paper. The research presented here was supported by awards to Jane Stuart-Smith from the Leverhulme Trust (F/179/AX), the AHRC, and the ESRC (R000239757), and to James M. Scobbie and Jane Stuart-Smith, from the ESRC, RES-000-22-2032 and RES-062-23-3246.

### *1.1. Derhoticisation in Scottish English in the twentieth century*

Scottish English is a range of varieties forming a sociolinguistic continuum between two poles, broad vernacular Scots spoken by working-class speakers at one end, deriving historically from Northern forms of the Anglian dialect of Old English, and Standard Scottish English (SSE), spoken by middle-class speakers at the other, continuing varieties of Southern English which were adopted by the upper classes from the seventeenth century onwards, and later used increasingly by middle-class speakers (e.g. Stuart-Smith 2003, Durand 2004). In the conurbations of the Central Belt of Scotland stretching between Edinburgh and Glasgow (Figure 1), home to most of the population, many speakers drift up and down the continuum according to formality, context and interlocutor (Aitken 1984). In these urban areas, stratification by social class is still strongly adhered to at both ends of the continuum, with a continual process of social (and geographical) mobility in between (e.g. MacFarlane & Stuart-Smith 2012).

INSERT FIGURE 1 HERE

Figure 1. The Central Belt of Scotland (see inset) showing the cities of Glasgow on the west, Edinburgh on the East, and Livingston in between (from Lawson et al. 2008).

Accents of English which have a phonological specification of consonantal /r/ in coda position (also called ‘postvocalic /r/’) in words such as *car*, *card*, are often referred to as ‘rhotic’. Scottish English is the classic rhotic variety of English in the UK (Wells 1982). Although /r/ was once an apical tap [ɾ] and often a trill [r] (Grant 1914, Johnston 1997), at least since the turn of the nineteenth century, derhoticisation in working-class speech, alongside an increasing use of approximant forms of /r/, have led to a sociophonetic continuum in the realization of postvocalic /r/. By derhoticisation, we mean either, diachronically, the gradient phonetic lenition process from trill towards a complete loss of /r/, or, synchronically, productions of /r/ weakly exhibiting few or none of the correlates typically attributed to its rhotic status. We survey the evidence for derhoticisation briefly below.

Reports of weak rhoticity in the realization of postvocalic /r/ date back to the early twentieth century, when reports of accent variation are first available. They relate to Scottish English spoken on the West coast, and specifically, as characteristic of the urban speech of the ‘degenerate Glasgow-Irish’, to whom numerous undesirable speech and language habits were attributed, including the infamous glottal stop (Trotter 1901 in Johnston 1997: 511). Polite speakers were noted to use the apical trill [r] or tap [ɾ] (Williams 1909, Grant 1914), or the postalveolar approximant [ɹ] (though, at this point, approximant /r/ was not considered a ‘Scottish sound’ by Grant and Dixon 1921, in Romaine 1978). All these realizations are attested in the very short reading passages recorded by William Dögen for the Berliner Lautarchiv in 1916/17 from young male speakers from Glasgow and surrounding areas (Richmond 2013). By 1938, approximant [ɹ] was a recommended realization for the ‘student of good speech’, as acceptable as [r], and more so if speakers wished to achieve the socially more desirable merger of /ʌ ɪ ɛ/ to /ɜ/ in a prerhotic context, e.g. in the words *fur*, *first* and *herb* (McAllister 1938; Lawson et al. 2013, forthcoming).

The earliest indication of derhoticisation in Edinburgh is indirect, from observations made in the Edinburgh Articulation Test (EAT), a standardized study of articulation in children’s speech aged 3.0 to 5.6 carried out in the late 1960s. The authors of the EAT coded vocalized variants along with consonantal /r/, stating: “many Scottish 2½-year-old children used a diphthong in positions where they later developed one of the many forms of [r]. *As this diphthong may also be an acceptable adult realisation, it had to be considered correct in this context.*” (Anthony et al. 1971: 6, in Scobbie et al. 2007, their emphasis). Note that such diphthongs may also have been picking up recessive upper-middle-class non-rhoticity.

A clearer picture of derhoticisation in Edinburgh is made possible thanks to two early sociolinguistic studies, carried out by Romaine (1978) and Johnston (Speitel & Johnston 1983, Johnston 1985). Romaine’s study concentrated on working-class children. Her results showed that boys were less rhotic than the girls, who also used more instances of postalveolar [ɹ], as opposed to tapped or trilled variants. Non-rhoticity was also more

common in the wordlists than in spontaneous speech. Romaine interpreted the non-rhoticity in the boys as a vernacular change from below taking place in Scots, “which happens to coincide with a much larger national norm” (i.e. ‘national’ in a UK sense, indicating non-rhoticity in RP, p.155). She saw non-rhoticity as carrying covert prestige, and part of a local system of differentiation from the more socially-desirable postalveolar approximant [ɹ] favoured by the girls, associated with middle-class speakers and prestigious varieties of Highland English (p. 156).

Johnston’s study worked with a much larger socially-stratified corpus of adults. He observed two very different kinds of non-rhoticity: that found in older (55-79 year old) Upper Middle-Class women, and that at the opposite end of the social-gender continuum, lower working-class men (18-55 years old), who showed vocalization to a ‘strongly pharyngealized vowel’. Such an outcome is not surprising since articulated /r/ in this speaker group is typically ‘dark’, with secondary pharyngealization. Johnston also found that in coda position, postalveolar [ɹ] was favoured particularly by younger female speakers, and in more formal styles. He suggested that postalveolar [ɹ] was “a recent innovation, probably from middle-class RP, into Edinburgh speech” (p. 27). Johnston interpreted the motivations for both changes in terms of the social dynamics within Scotland. Derhoticisation was identified as showing ‘street-smart’ associations; rhoticity in the middle classes was seen as reflecting constructions of a resurgence of Scottish identity in the Scottish middle-classes, expressed in a ‘home-grown model of Standard Scottish English’ used in preference to, and a reaction against, earlier local Scottish prestige models close to RP.

Back on the West Coast, Macafee’s (1983: 32) description of Glaswegian dialect, outlined similar derhoticisation to plain or pharygealised vowels in working-class speakers. Subsequent quantitative analysis of a socially-stratified corpus of Glaswegian collected in 1997 confirmed substantial derhoticisation in working-class speakers, especially adolescents (Stuart-Smith 2003; Stuart-Smith et al. 2007). Derhoticised reflexes fell into two main categories: pharygealised/uvularised vowels, favoured by boys in a specific phonological context (before a consonant, e.g. *card*); and plain vowels with no audible secondary ‘colouring’, favoured by girls in unstressed prepausal position, e.g.

*better#*, though both groups showed numerous instances of both variants. Middle-class speakers tended to be rhotic, with both older and younger speakers favouring postalveolar and/or retroflex approximants, especially younger middle-class girls. (If articulatory /r/ was produced by working-class speakers, it was usually a tap.)

Overall, the evidence for the twentieth century suggests the development of a socially-stratified rhotic-derhotic continuum in the Scottish English of the Central Belt, with weakly articulated, or vocalized, rhotics in working-class speech contrasting with audibly strong rhotic approximants in the aspiring middle-classes. We now turn to the sociolinguistic evidence for the progress of derhoticisation, and the corresponding development of the continuum, in the early 21<sup>st</sup> century.

## 2. *Derhoticisation in Scottish English in the 2000s*

In 2003, a further corpus of Glaswegian was collected from an age-stratified sample of working-class speakers from the same area as the 1997 corpus (e.g. Stuart-Smith 2006; Stuart-Smith & Timmins 2010). Figure 2 shows the substantial derhoticisation that was found in these speakers. Like Romaine (1978), derhoticisation was more prevalent in read wordlists. This stylistic shift away from the regional standard norm (rhoticity) in a reading task confirms that this feature still carries the kind of covert prestige suggested by Johnston.

### FIGURE 2 ABOUT HERE PLEASE

Figure 2. Distribution of variants of postvocalic /r/ in 48 speakers of Glaswegian in 2003, n = 1889. M = male, F = female; 1 = 10-11 years; 2 = 12-13 years; 3 = 14-15 years; 4 = 40-60 years. [r] = articulated variants of /r/; [V^] = vowels with audible pharyngealisation/uvularisation; [V] = plain vowel; [Vh] = vowel followed by audible frication.

Only six years had elapsed by the time we collected the 2003 corpus, so it is difficult to know the extent to which variation over this time reflects real-time change (Labov 1994). Comparison of the percentage of use of the plain vowel variant for coda /r/ for individual speakers in 1997 (8 children) with those recorded in 2003 (36 children)

suggests that derhoticisation is a very gradual change in progress. The speakers from 1997, shown as dark bars, fit within the distribution of the speakers from 2003; see Figure 3.

#### FIGURE 3 ABOUT HERE PLEASE

Figure 3. Percentage of the plain vowel variant for coda /r/ used by 42 speakers, 36 recorded in 2003 (pale bars) and 8 recorded in 1997 (dark bars). The left chart shows female speakers, the right male speakers.

Previous studies had concentrated on the two cities at either end of the Central Belt. In 2007 a corpus of speech and articulatory data (tongue movement) was collected from working-class adolescents in Livingston, a new town, in between, but lying closer to Edinburgh than Glasgow (Figure 1); Lawson et al. (2008). Auditory transcription showed some derhoticisation, but on average only 20% of all postvocalic /r/, which is considerably less than the amount found in Glasgow. Also unlike Glasgow, the most common environment for non-rhotic tokens was in stressed syllables in utterance final position (e.g. *car##*), though the next most likely context was in unstressed syllables in utterance final position (e.g. *better#*).

#### FIGURE 4 ABOUT HERE PLEASE

Figure 4. Bar graph showing the percentage of auditory variants used by each socioeconomic and gender group in the ECB08 corpus. WC/MC = working/middle-class; M/F = male/female. Paler grey segments represent rless and weakly rhotic variants, while darker grey segments represent strongly rhotic variants.

N=139. From Lawson et al. (2011), Fig. 2.

A year later, a socially-stratified audio and articulatory corpus (ECB08) was collected from middle-class adolescents in Edinburgh, and working-class adolescents again from Livingston. The study was designed to further explore possible articulatory mechanisms for derhoticisation. The auditory assessment of postvocalic /r/ drawn from the wordlist confirms more weakly-articulated /r/ and derhoticisation (pale grey segments) in working-class speakers towards the East, and illustrates well how the rhotic-derhotic continuum is constrained by social class and gender (Figure 4; Lawson et al. 2011).

It is clear that these recent data continue the earlier trends. Middle-class, and especially female, speakers are leading a change from above towards audibly ‘strong’ approximant /r/. These changes exploiting the variant [ɹ], which may be of Anglo-English origin, to mark both more confidence in a specifically *Scottish* (not UK) middle-class identity (Johnston 1985), and social differentiation from Scottish working-class identities (Douglas 2009). Working-class speakers on the other hand are participating in long-term vernacular change from below, resulting perhaps in the completion of derhoticisation which will be non-rhoticity. The earliest reports pin the latter change to the turn of the twentieth century, but the change may have started much earlier. The progress of derhoticisation varies according to location, but is more advanced in the more populous western conurbation.

Another important aspect of Scottish derhoticisation is how it relates to non-rhoticity in English English. For it cannot be ignored that in some phonetic contexts, e.g. following /a/, the derhoticised reflexes in Glasgow appear strikingly non-rhotic, making the outcome phonetically very similar to the non-rhoticity found in the UK standard (and indeed non-standard) varieties of English English (Romaine 1978). Moreover, the recent large-scale study of rhoticity along the Scottish-English Border has also found derhoticisation in younger speakers, though with significantly more at the western end (Gretna) than in the more Scottish, east-coast, town of Eyemouth, which aligns with attitudes of Scottishness (Llamas 2010). Pukli and Jauriberry (2011) also report some derhoticisation in the rural south-western city of Ayr, as well as the substantial appearance of postalveolar [ɹ] in onset position, and more generally in young female speakers. Just as other consonantal changes appear to be making their way north (e.g. Th-fronting, L-vocalisation; Stuart-Smith et al. 2007), there is a possibility that the Glaswegian non-rhotic outcome could also reflect the effective confluence of two streams of change, one a vernacular change within Scots, and the other a contact-induced change from non-rhotic varieties of English. In order to consider the empirical evidence for this, in the following section we put derhoticisation in the context of the wider system of



changes in progress in Glaswegian, and the social factors which are involved in their transmission.

### ***3. Social factors in Glaswegian derhoticisation***

The most recent study of derhoticisation of /r/ in Glasgow was undertaken as part of a broader variationist project. Its aim was to consider the role of a large range of social factors in several sound changes in progress in Glaswegian, including opportunity for contact with speakers of dialects furth of the city, and the possible influence of the broadcast media. Also in the 1997 corpus, derhoticisation of postvocalic /r/ was found in the speech of those working-class adolescents who were leading in the rapid adoption of some consonant features typically associated with London and southern English, specifically the use of [f] and [v] for /θ/ and /ð/ (TH/DH-fronting), and vocalization of coda /l/ to a high back (un)rounded vowel (L-vocalisation). That these speakers were also the least geographically and socially mobile posed a challenge for contact-based theories of the diffusion of these changes (e.g. Trudgill 1986), and the media themselves suggested that watching television, and in particular, dramas set in London, like the exceptionally popular soap, *EastEnders*, as a key factor.

The Glasgow Media Project constituted the first comprehensive systematic sociolinguistic investigation of the influence of the broadcast media on language change, by focusing on the possible role of exposure to, and psychological engagement with, London-based TV dramas on Glaswegian vernacular phonology. Three groups of linguistic variables were considered:

- consonant innovations: e.g. TH-fronting. Three rapid changes in Glaswegian look like instances of diffusion from Southern varieties of English English, which took off in the 1990s, though they are sporadically reported in Scottish English much earlier (Macafee 1983, Anthony et al. 1971);

- ongoing vernacular changes: e.g. derhoticisation of postvocalic /r/. As noted before, this change appears to be system-internal, though the final outcome (e.g. non-rhoticity) can coincide phonetically with English English norms;

- more stable sociolinguistic variation: e.g. realization of the vowels /a/, /u/ and /ɪ/.<sup>2</sup>

Only the consonant innovations have been explicitly linked with exposure to London English on the television. However to test the hypothesis that television might be a contributory factor in the innovative changes, we needed also to test those variables for which media influence has never been mooted, and so vowels and derhoticisation were included in the study.

The auditory variants for the consonant innovations (e.g. [f], [v]), and derhoticisation (/Vr/ sequences realized as a plain vowel with no velar or pharyngeal quality), and F1 and F2 of /a u ɪ/, for read (wordlists) and spontaneous (conversational) speech were the dependent variables in a series of regression models constructed for the 36 adolescent informants. The independent variables consisted of representative linguistic factors (e.g. position in the word, adjacent phonetic context), and a large array of extralinguistic factors: opportunities for dialect contact with speakers of other English dialects; attitudes to dialects elicited from responses to audio recordings and paper surveys; engagement and participation in a range of social practices; preferences for music and radio, film (cinema, DVD, video); activity on the internet and engagement with computer games and computer-mediated communication; and exposure to, and psychological engagement with, the television. The variables were drawn from a structured questionnaire completed by each informant, an informal interview with the fieldworker, their own spontaneous speech recordings with their friends, and participant observation by the fieldworker during the period of data collection. Full details and results of the regression study can be found in Stuart-Smith et al. (2013, forthcoming).

The main findings were:

- The consonant innovations were strongly constrained by linguistic factors and by several extralinguistic factors including: participation in anti-school social practices, such as adopting Glasgow street style in place of school uniform; strong psychological and emotional engagement with the London-based TV soap opera,

---

<sup>2</sup> That was our hypothesis at the time. In fact the new Glasgow Real-Time Project is demonstrating real-time change in /u/ (e.g. Rathcke et al. 2012).

*EastEnders*; reported contact with friends and relatives in England; and more weakly, with positive attitudes to London accents.

- The vowel variables showed almost exclusively significant effects for linguistic factors, with very little evidence for social factors of any kind.

- o The results for derhoticisation of postvocalic /r/ were split according to speech style. (i) In spontaneous speech derhoticisation patterned like the vowel variables: the predominant effects were for the linguistic factors, with very little evidence for social factors. (ii) In read speech derhoticisation showed a similar pattern to the consonant innovations: both linguistic and social factors were significantly correlated. Plain vowels for /Vr/ were more likely in unstressed prepausal position (e.g. *better#*), and they were linked with anti-school social practices, strong psychological engagement with *EastEnders*, and the ability to correctly identify their own local accent from a recording, amongst other factors (including participation in sport and playing football); only dialect contact proved to be consistently non significant.

To summarize: there is no evidence that direct contact with non-rhotic English speakers promotes derhoticisation. But indirect contact with non-rhotic London English, by psychologically engaging with a TV drama set in London, does seem to be a factor, but only for a particular speech style, reading a list of words out loud.

The results for derhoticisation indicate that the change is not entirely driven by system internal forces. At the same time, they contribute to our understanding of media influence on speech more generally. The evidence from Glasgow shows that only some phonological features are linked to engagement with the television. This supports an extension of existing models of media influence in mass communications theory to language, specifically that speaker/viewers use their social and linguistic knowledge to ‘decode’ televised speech, so here Glaswegians parse *EastEnders* through the filter of their own experiences as active members of actual speech communities within the city (Hall 1980; Gunter 2000; Stuart-Smith 2011). The assumption is that viewers largely filter out aspects of media language which are irrelevant in terms of social meaning and linguistic structure (which is probably the majority of most experienced media material). But sometimes a viewer’s existing features may be enhanced provided that there are points of reciprocity and alignment with the viewer’s own local social context and

linguistic system (which is probably quite rare). So the consonant innovations look like diffusing features ‘from outside’ the dialect, hopping north from London. While there is some support for dialect contact being involved, closer up they look fundamentally like local system-internal variation which is, as it were, bubbling up, developing a variety of social symbolic functions, which in turn speed up the changes in progress (Eckert 2000). Media influence represents an additional factor through which speakers enhance their existing variation, thus fuelling their rapid acceleration through the system and the community.

Derhoticisation has been underway for many decades in Scottish English, apparently without influence of English English. Only in read speech is derhoticisation linked with indirect contact with London English via the TV. This helps unpick the processes of media influence further. When we recorded the working-class adolescents reading the wordlists, rather than read them ‘correctly’ (i.e. approximate Standard Scottish English, e.g. Labov 1972), our informants produced distinctly non-standard variants. Overall a specific position, or stance towards the task and fieldworker was taken (Jaffe 2009), as if distancing themselves and their speech from the University. The wordlists were rattled off, punctuated with laughter; they were highly performative, in the sense of Baumann’s construction for an audience (Coupland 2007). In terms of variation, the wordlists showed increased use of consonant innovations, and more derhoticisation. Previous research on stance-taking through language has noted that media representations can simplify social-indexical relationships, and so speed up linguistic appropriation from the media (see e.g. the spread of the catchphrase ‘Whassup?’ in American English; Bucholtz 2009: 288). Aspects of language which index nuances of interpersonal interaction, and subsequently local micro-social relationships, can then be used in the media, e.g. in advertising, with much broader indexical referents.

We hypothesize that stance, and/or other kinds of social informativity of linguistic variation (Pierrehumbert 2006, Eckert 2008), may be a determining factor in whether speaker/viewers’ sociolinguistic systems may respond to media language. The enhancement of viewers’ existing features may depend on the implicit recognition, or mapping, of linguistic features indexing particular stances in media language, with the possible indexing of stancetaking in their own interactions. Crucially, being perceivers

and producers of social variation, or being listeners using their ‘speaking brain’ (Keith Johnson personal communication), is also important here; Kuhl, e.g. (2010). The interesting point about the link between engagement with the TV and derhoticisation is of course that this change has never been interpreted as a contact-induced change. These results emphasize the importance of the speaker/viewer’s local social-phonological system in the decoding of televised speech. They also suggest mechanisms for how existing local variation could become accelerated through indirect contact with accent features, albeit through strong psychological and emotional engagement with a television programme and its characters. We suspect that direct contact with English English does not emerge as a factor precisely because this is mediated by ideological and attitudinal factors relating to nationality and non-rhoticity (Llamas et al. 2009).

Thus, teasing apart the social factors that contribute to the progress of derhoticisation is both informative in understanding the change itself, and for modeling media influence on speech. There is indeed some evidence to support the view that non-rhoticity in western Central Scottish English reflects the outcome of two streams of change, though the nature of the contact-induced change needs to be refined to indirect contact with non-standard English via the broadcast media. But we still need to discover the phonetic mechanisms underpinning derhoticisation, and the rhotic-derhotic continuum; in order to do this we must consider the phonetic data – and how they might be represented.

#### ***4. Scottish derhoticisation and the listener***

The variation observed in the Scottish English rhotic-derhotic continuum, provokes two challenging questions: (1) What is the phonetic nature of the derhoticised reflexes? (2) How can we best capture this complexity? Until recently representing sociophonetic variation was limited to characterizing aspects of the recorded speech signal, by auditory or acoustic analysis. Whilst it is increasingly assumed that acoustic analyses are superior to auditory ones, and certainly they have the advantage of yielding continuous measures which are amenable to more robust statistical analyses (e.g. Warren

& Hay 2012), both are equally valid. Each gives a different (and incomplete) picture of the ‘same’ thing; both are connected, but not in straightforward ways, and in turn make inferences about underlying articulatory gestures. In this, and the next two sections we review previous and ongoing phonetic work on derhoticisation which exemplifies these points. We begin by considering the view from the listener, both the analyst and the speech community. In section 5 we shift perspective to look at acoustic representations. In section 6 we move closer to articulation, using Ultrasound Tongue Imaging.

#### *4.1. The listener as analyst: auditory phonetic representations of derhoticisation*

All the studies discussed above used impressionistic or auditory transcription. Using this method, analysts categorise the auditory continuum of variation in ‘articulatory’ terms, i.e. the analyst constructs a kinaesthetic interpretation of the possible articulatory strategy used by the speaker, and then represent it using IPA symbols (Ogden 2009). Transcription can be more or less detailed, but usually results in fairly broad?, discrete categories, which make strong assumptions about the articulatory gestures underlying the auditory objects. Whilst auditory transcription is a valid and useful method of representing phonetic variation, we need to be mindful that it yields auditory, not articulatory, objects. It also requires the analyst to broadly divide up and assign parts of the auditory continuum to one or other categories, whereas listeners may feel that aspects of more than one category may be involved. Social-indexical ine-grained variation may not be easily audible even to trained phoneticians (Docherty & Foulkes 1999).

Each group transcribing derhoticisation came up with different solutions,<sup>3</sup> which in turn coloured their theoretical perspective. For example, recognizing many possible

---

<sup>3</sup> Speitel & Johnston (1983) and Stuart-Smith (e.g. 2003) used very narrow auditory phonetic transcription and identified a range of different kinds of derhoticised and/or vocalic outcome, which can be represented either as extremely weak uvular approximants, or vowels with secondary pharygealisation. Romaine recognized this phonetic complexity but opted to represent a simplified set of categories, grouping plain and coloured vowels together as complete deletion. Lawson et al 2008 simply divided variants into rhotic and non-rhotic.

variants emphasizes gradient progression of the change, as opposed to coding with or without final /r/, which points to the final outcome (contrast ‘derhoticisation’/‘R-vocalization’ with ‘R-Loss’). For all, the transcription of the derhoticised variation was extremely difficult, and this motivated a small-scale study to investigate this analytical task (Stuart-Smith 2007).

A subset of the 2003 Glasgow corpus was selected, 12 male working-class informants, nine adolescents, with three from each age group, and three adults. All the adolescents were observed to show derhoticisation in the main study. A subset of words were selected from the larger wordlist, in which /r/ follows the low vowel /a/: *heart, barn, farm, car, far, card*. These were subjected to a narrow auditory phonetic transcription by three phonetically trained transcribers: 1: CT, a Scottish-English, rhotic middle-class speaker from Edinburgh; 2: JSS, an English-English, non-rhotic middle-class speaker from Southern England; 3: RL, a Scottish-English, rhotic middle-class speaker from a small town just south of Glasgow. The results of the transcriptions are shown in Figure 5.

#### FIGURE 5 ABOUT HERE PLEASE

Figure 5. Results of the auditory transcription of postvocalic /r/ in word-list data read by 12 male Glaswegian working-class speakers, organised into four age groups (1= 10-11, 2= 12-13, 3=14-15, 4= 40-60). The judgments of the three transcribers (CT, JSS, RL) are shown in each chart from left to right. White = articulated /r/, spotted = pharyngealised/uvularised vowels, grey = plain vowels, striped = vowels followed by [h] or [ɦ], from Stuart-Smith (2007: 1308).

The results are striking. Each transcriber hears the same signal, but transcribes and categorises it differently from each other (see also Plug & Ogden 2003). All heard some derhoticisation, CT the least, and RL the most – so interestingly the outcome is not straightforwardly predicated on the transcriber being rhotic (Yaeger-Dror et al. 2009), though perhaps differential experience of the rhotic-derhotic continuum, and/or the socially symbolic nature of derhoticised variants might play a role. Recall that derhoticisation is more advanced on the West than the East. We also found that whilst transcribers effectively segmented the auditory continuum at different points, they were internally consistent.

There is also another key shared feature. All three transcribers found that they could not assign what they heard only to two categories, ‘plain vowel’ or some kind of articulated /r/ (the phonetic variants are grouped together for this representation but ranged from weak approximants to weak taps). A third auditory category was needed for variants which fell between articulated /r/ and no audible articulation at all, which could be termed either as ‘extremely weak uvular approximants’ or as ‘pharyngealized or uvularized vowels’. This could be interpreted in a prescriptive way as analysts simply being unable to implement the IPA categories appropriately. But we will see that the acoustic, and especially the articulatory, data show that a category to accommodate such a variable percept – hearing sometimes a consonantal gesture and sometimes not – is well motivated.

#### *4.2. The listener in the community: evidence from speech perception*

An alternative view to that of the analyst can be drawn from perceptual evidence from the community – how listeners parse, and/or respond to variants along the rhotic-derhotic continuum. Carey (2010) carried out a small-scale study of cross-linguistic dialect perception, looking at Glaswegian and Southern British English (SBE) listeners’ responses to stimuli from both dialects. Judges listened to three pairs of sentences which varied according to whether postvocalic /r/ was present or absent, e.g. *That surprise for the child* vs *That’s a prize for the child*, or *The congregation certainly likes arms* vs *The congregation certainly likes psalms*, and then had to write down what they heard (the stimuli examined a large number of phonological differences between SBE and Glaswegian). Glaswegian listeners found it as difficult as SBE listeners to ‘recover’ postvocalic /r/ in such sequences, even in the stressed monosyllable *arms*.

MacFarlane & Stuart-Smith (2012)’s matched guise study considered social evaluation. The same talker produced recordings of pairs of words which varied in the realization of a single variable. Listeners were led to believe that two speakers, Lee (‘regular Glasgow’) and Phil (‘socially-aspirational Glasgow Uni(versity)’) had produced the recordings, and were given only a group of brand logos for each ‘speaker’ as their guide to the lifestyles of the ‘two’ men. Three out of the four experimental variables



related to /r/. The realization of onset /r/; the duration and quality of the final syllable of disyllabic words such as *number* (longer for Glasgow Uni, shorter and less rhotic for regular Glasgow); and the quality of the prerhotic vowel in words like *nerve* and *pearl* ([ə] is associated with Glasgow Uni – and also with vocalic rhoticisation; [ɛ] is associated with a following tapped /r/ variant and Regular Glasgow). Listeners were very good at correctly socially categorizing the ‘talkers’ using the *number* and *nerve* variables, i.e. the two variables which related to realization of postvocalic /r/. But the realization of onset /r/ was only categorized at chance level, refuting the hypothesis that taps in this position associate more with ‘regular Glaswegian’ speech, that is, working-class Glaswegian speech.<sup>4</sup>

These two studies both show that derhoticisation is also taking place perceptually for members of the community, and is not only restricted to the domain of the analyst. Both ends of the rhotic-derhotic continuum also still seem to carry the kind of locally-salient social meanings that were proposed by Johnston for Edinburgh. But if we want to pin down what listeners are responding to, it is clear that we need to go further than the admittedly tricky auditory categorization. Our next attempt was acoustic analysis.

### 5. *The acoustic characteristics of derhoticising /r/*

The difficulties with auditory percepts which were challenging to auditorily categorise, and themselves variable, motivated an acoustic analysis of the data whose

---

<sup>4</sup> This last result is intriguing since it suggests that the realization of coda /r/ carries more meaning for these speakers, than that of onset /r/. If this is right, this might also account for Johnston’s suggestion that postalveolar approximant /r/ spread from English English into Edinburgh English in onset position. Pukli & Jauriberry’s (2011: 88) findings from Ayr that onset /r/ is increasingly being realized by an alveolar approximant [ɹ] in Ayr are also congruent. So too are the similar shifts observed at the western end of the Scottish/English border by Llamas 2010. The originally ‘English’ variant may have slipped more easily into the array of /r/ variation in this environment, becoming Scottish, but unmarked as such, precisely because variation in onset /r/ does less social ‘work’ than coda /r/. Our current work on articulation of /r/ is interrogating this assumption further.

auditory transcription was discussed above in §4.1 (Stuart-Smith 2007). Since it was also unclear whether the final outcome of derhoticisation to plain vowels is leading to a merger (recall that weakened /r/ is now perceptually variable even to Scottish listeners, §4.2), we included minimal pairs. To recap, we considered the acoustics of coda /r/ in the following words: *heart/hat*, *barn/ban*, *farm/fan* and *car, far, card*, in 12 working-class speakers, nine boys and three men.

We carried out a qualitative visual analysis of the spectrograms, and then used a parametric analysis of acoustic properties of the syllable rime, so e.g. *c-ar*, following the successful application of this method to variation in postvocalic /r/ in Dutch (Plug & Odgen 2003). This also addresses the practical difficulties of segmenting final /r/ consonants which were effectively no longer there. Using Praat, we labelled the waveform for the beginning and end of the vocalic portion (i.e. the entire duration of the vowel+/r/ portion of the syllable rime), and then measured the duration of the vocalic portion, the vowel quality in terms of the first three formants at the midpoint, and vowel tracks for the last five glottal pulses, again for the first three formants. Formant measures were extracted using Praat, and then corrected by hand.

The classic acoustic ‘signature’ of approximant /r/s, and also some trills and taps, is a lowered third formant (Lawson et al. 2011a; though see Heselwood & Plug 2011). The lowered F3 relates to the dimensions of a large cavity in the front of the vocal tract arising from specific articulatory gestures. The rather different configuration for uvular /r/ shows a different pattern of high and/or raised F3. Visual inspection of the spectrograms provided the following acoustic information for the four auditory variant categories shown above in Figure 5:

- articulated /r/: This included taps, a few weak approximants, and a single trill in the oldest man. The taps showed the expected momentary reduction in amplitude across the frequency range (Figure 6a), and the trill had four such dips visible, reflecting four short interruptions in airflow (Figure 6b). In the few tokens of /r/ which were heard as (weakly) articulated approximant /r/, it is just possible to see the faint trace of the third formant dropping towards the end of the word, though just as striking is the reduction of amplitude above F2 (see Figure 6c).

The other three variant categories capture different stages of audible erosion of the rhotic consonant:

- pharyngealized/uvularized vowel: These variants sound like extremely weak uvular approximants, or vowels with pharyngealization/uvularization. The primary acoustic characteristic is reduction in amplitude where /r/ would be expected. The weakened F3 is either flat or rising slightly (see Figure 6d);

- plain vowel: No primary or secondary articulation for a rhotic consonant was audible. The spectrograms typically show flat first and second formants, with very little energy above F2 (see Figure 6e). Inspection of successive spectra shows a very weak third formant which rises towards the end of the vocalic portion and into the voiceless period;

- vowel followed by audible frication: A small number of plain vowels sounded as if they were followed by a very weak fricative, possibly glottal, pharyngeal or even uvular. In Figure 6f, the vowel gives way to a period of very weak energy, with initial energy loss in F3, and then voicing ceases, though a period of very weak aperiodic noise is still visible for several ms.

#### INSERT FIGURE 6 ABOUT HERE PLEASE

- |  |  |
|--|--|
| (a) <i>farm</i> with tap (adult male)                | (d) <i>card</i> with pharyngealized/uvularized vowel (12 yr-old boy) |
| (b) <i>car</i> with trill (adult male)               | (e) <i>car</i> with plain vowel (11 yr-old boy)                      |
| (c) <i>far</i> with weak approximant (14 yr-old boy) | (f) <i>far</i> with vowel followed by weak frication (14 yr-old boy) |

Figure 6. Spectrograms illustrating the four auditory variant categories shown in Figure 5. Articulated /r/ is shown on the left in (a)-(c); vowel variants on the right – pharyngealized/uvularized vowel (d), plain vowel (e), and vowel followed by weak frication (f). All recordings were made in 2003 in Glasgow.

Neither first and second formant measures, nor durations, differed according to whether an articulated /r/ was audibly present or absent. Derhoticisation is not reliably distinguished through these measures. On the whole F3 was very difficult to measure because – as was observed – towards the end of the vocalic portion, where an acoustic reflection of the /r/ sound might be expected to be seen, there was a sharp drop in intensity in and above the region of F2, and in the F3 region. If it was possible to pick out

F3 in speakers whose variants were audibly less rhotic, F3 was either flat or rising slightly, consistent with uvularization. This is illustrated in a comparison of the formant tracks from the most audibly rhotic boy with his much less rhotic-sounding friend (see Figure 7). A further result is that the derhoticized outcomes of /r/, even plain vowels, are still significantly distinct from words without <r>, so e.g. derhotic *heart* shows a longer, more retracted vowel than *hat*. This suggests that, at least for wordlist data, there is not yet a loss of phonological /r/, which is hinted at by Carey's (2010) results; it is likely that as in other non-rhotic varieties of English, the contrast will be maintained by differences in the vowel system (for further discussion of the impact of rhoticity on Scottish vowels, see Lawson et al. 2013).

#### FIGURE 7 ABOUT HERE PLEASE

Figure 7. Handcorrected time-normalized formant tracks taken at the end of the vocalic portion and for each of the five preceding pulses, for the first three formants for two speakers: a) 14 year-old boy heard as rhotic, shows slight dip in a high F3 in most words with /r/ (this boy produced *far*, Figure 6c). b) 14 year-old boy heard with mainly pharyngealized vowels for words with /r/, shows high, flat or rising F3, with weak amplitude (this boy produced *far*, Figure 6f).

The outcome of the acoustic analysis is not as helpful as we had hoped. In part this is because the reflexes of derhoticisation do not relate easily to known acoustic parameters. Rather the clearest common characteristic is a reduction of acoustic energy above F2. On the one hand, these stretches of very weak formant energy, with and without, voicing, may help account for the variable auditory percepts of rhoticity. That is to say, sometimes there is, and sometimes there is not, some kind of secondary pharyngeal articulation, and so the residue of an articulated /r/ is still present. But on the other, lack of energy in a specific frequency region makes it difficult to identify and measure formants in and above that region. Quantitatively capturing such acoustic weakening itself is also far from straightforward. This reminds us that acoustic analysis may not always be superior to auditory analysis; it is necessarily partial and more subjective than it might appear (Ogden 2009: 36). Thus the acoustic analysis moves us forward, but it still leaves us with another picture of the data, as opposed to a better

understanding of the mechanism of derhoticisation.<sup>5</sup> For this we need to turn to articulatory views of the phenomenon.

## ***6. Investigating derhoticisation using articulatory data***

Auditory-acoustic challenges led us to consider a different kind of phonetic representation, closer to the articulatory strategies involved, achieved using Ultrasound Tongue Imaging (UTI), and arising from a 2004 study of Dutch /r/ (Scobbie & Sebregts 2011). Our Scottish work is in progress, and in this section we report some key relevant findings from three recent studies carried out on the Eastern Central Belt, a small pilot reported in Scobbie (2007), a sub-project to assess the feasibility of UTI for sociolinguistic fieldwork (WL07 corpus from Livingston; Lawson et al. 2008), and a socially-stratified articulatory speech corpus, with middle-class speakers from Edinburgh and working-class speakers from Livingston (ECB08 corpus; e.g. Lawson et al. 2011). Initial results from Glasgow are reported in Lawson et al. (2013, forthcoming) Full details of our UTI set-up, the methods for each study, and full analytical results are given in each of the references cited. After brief comments on the technique itself, we show how UTI reveals a probable cause for both the auditory, and the acoustic ambiguities presented by derhoticisation, as well as an articulatory basis for the socially-stratified rhotic-derhotic continuum, in terms of gestural timing (§6.1), tongue configuration (§6.2) and the extent to which these can be accessed (or not) by the listener (§6.3).

UTI makes use of ultrasound technology designed for usual medical research, capturing analogue video showing visual dynamic representations of tongue

---

<sup>5</sup> More may be learnt from a psychoacoustic representation than an acoustic one, given Heselwood & Plug's (2011) recent experiments which strongly suggest that the key perceptual feature of rhoticity (typical of approximants) may be "not a low-frequency F3 *per se*, but rather a single perceptual formant in the F2 region, which we might label F-rho" (p. 870). Lennon's (2011) application of a Bark difference metric (Z3-Z2) to the real-time increase in strong rhotics in middle-class speakers in Glasgow's northern suburbs, suggests that this could be a useful analytical tool for future research.

configuration and tongue movement, usually, but not exclusively, in sagittal orientation (see Figure 8).

#### FIGURE 8 ABOUT HERE PLEASE

Figure 8. Midsagittal image of the tongue surface produced using a Concept M6 medical ultrasound machine. The tongue root is to the left of the image and the tip is to the right of the image.

In our setup the ultrasound probe is held under the chin by a stabilising headset, and the screen displays a 2D fan-shaped image, showing the water-air interface, i.e. the tongue surface, as a bright white line, thanks to the great intensity of reflections of ultrasound pulses back to the probe. To some extent the internal muscle structure of the tongue can also be seen. It is possible to visualize almost the whole of the mid-sagittal shape and location of the tongue, root, dorsum, front, and sometimes the tongue tip – though the tongue tip is often not visible when it is raised, due to the presence of a sublingual airspace. We use specialist software, Articulate Assistant Advanced™, to capture, process and analyze the data (Articulate Instruments Ltd. 2011).

Whilst UTI gives instant dynamic and static impressions of tongue movement which are immediately informative, quantifying UTI data is challenging and techniques are still under development. Data are also less direct than it might appear, both because of the basic video frame rate (only 30 frames/sec), and the way in which images are constructed by video-output ultrasound machines. This means that ultrasound data are somewhat removed from actual articulation, being both partial and processed. Nevertheless, UTI offers sociophoneticians an excellent tool for investigating speech articulation, both because it is safe and non-invasive, and because – despite the visible headset and need for technical personnel – the method can have minimal quantitative impacts on speech style. Lawson et al. (2008) shows that in fact stylistic variation is more dependent on the speakers' relationships with their interlocutors, and the presence of friends and peers, than the physical context induced by the equipment. Unlike speakers faced with just a microphone for speech recording, articulatory participants can also be ethically misdirected through a focus on the fact that the recordings are designed to record “changes in the shape of the tongue”, which incidentally requires speech.

### *6.1. Derhoticisation and gestural timing*

The UTI data from the pilot data and WL07 corpus uncovered a possible mechanism for derhoticisation in terms of gestural asynchrony. Recall that the auditory transcription was challenging because of the variable percept of sometimes hearing a consonantal gesture and sometimes not, but also from strong pharyngealisation on the vowel.

The articulatory data suggest that derhoticised postvocalic /r/ in our Scottish speech samples involves both (1) an early tongue root retraction gesture and (2) a delayed tongue tip raising gesture, though a systematic study remains to be carried out. An early tongue root retraction gesture could account for the modification of prerhotic vowels, specifically retraction and pharyngealisation of these vowels. The delayed tongue-tip raising gesture means that the maximum of the /r/ gesture is often masked by following consonants, or, prepausally, can occur after the offset of voicing, leaving the /r/ partially or completely inaudible.

This timing, weakening and interarticulatory dissociation of gestures may also account for the weakening of the amplitude of formant energy above F2 observed in the acoustic data. (Exactly how this is achieved is not yet clear, but it seems likely from Stevens' 1998 modeling of the acoustic consequences of the resonating cavities during the production of /r/ and /l/, that the shifts in gestures that we are witnessing are resulting in the formation of an additional cavity with strong damping properties on the spectrum, even before voicing has stopped.) In some speakers, faint dipping of F3 can be seen in a weakly noisy period after voicing has ceased, but this is not always easy to discern and timing of the covert tongue-raising gesture is variable. For example, in Figure 9, the tongue tip only starts to raise in frame 3, just as voicing is ceasing, and then continues to raise during the period of frication; the maximum raising in frame 6 occurs some time after voicing has stopped.

FIGURE 9 ABOUT HERE PLEASE

Figure 9. Key UTI frames of an adult male speaker from West Lothian, saying *car* showing a covert tip-raising gesture in the production of coda /r/. The ultrasound images correspond to the time point of the spectrogram. Moving through the frames, it is clear that the tongue front and tip begins to rise after voicing has ceased, and achieves its maximum raising well after.

Thus UTI shows how the timing of two of the gestures contributing to /r/, and in particular their relation to the offset of voicing, means that the primary anterior gesture for the rhotic cannot be reflected in the expected pattern of formant transitions during periodicity. Temporally drifting gestures would also explain the gradient loss of rhoticity. This, and the corresponding shifts in the resonating cavities, help explain the acoustic patterns observed for derhoticizing variants (§5; Figure 6d-f). It is also not surprising that the secondary pharyngealization becomes more audible – the tongue-root gesture is early – and that the /r/ is variably present – the tongue-tip gesture does occur, it just occurs much later, when voicing has stopped.

This account looks at postvocalic /r/ in a particular context, investigated due to the previous researchers finding that the phonological environment which most favoured derhoticisation in Scottish English was in stressed, utterance-final position, usually accompanied by vowel breaking, as in e.g. *It's near here* [hiʌ(r)] (see Figure 10, for the distribution of non-rhotic variants; see Romaine 1979: 45, Speitel & Johnson 1983: 28, Lawson et al. 2008).

#### FIGURE 10 ABOUT HERE PLEASE

Figure 10. Percentage of (un)stressed tokens in utterance-final and non utterance-final position that were audibly nonrhotic. n=1248. From Lawson et al. (2008).

Figure 10 shows that the second most likely phonological context for derhoticisation was in unstressed syllables, especially in utterance final position, as in Glasgow. Again, this may also relate at least in part to the kind of gestural asynchrony we described above as syllable lengthening is common in utterance-final position, allowing greater gestural asynchrony/dissociation, see Sproat & Fujimura (1993), or possibly also to gestural undershoot, assuming that speakers are likely to be producing an articulated /r/



with a tongue tip gesture, as for e.g. an apical tap. (We have no direct evidence from this particular set of ultrasound data because taps are too fast for the slow frame rate we used). The gradual loss of rhoticity in the history of English English also appears to have started in unstressed syllables (Dobson 1957), and even middle-class speakers who might otherwise be deemed thoroughly rhotic also show audible weakening in this position (Stuart-Smith 2003).

## 6.2. *Tongue configuration and derhoticisation*

Derhoticisation probably does not only arise from differences in timing, but also in tongue shape. In the Glasgow 1997 data, those who were likely to derhoticise were also more likely to use taps, if they showed articulated /r/, whereas more rhotic-sounding speakers used more approximants, especially auditorily-strong rhotics which we transcribed as retroflex approximants [ɻ] (Stuart-Smith 2003). Lawson et al. (2011) carried out a further investigation using the eastern Central Belt ECB08 corpus. The design consisted of two parallel analyses of the same data.

The first was an audio-rating analysis of randomized tokens, carried out using the independent classification of tokens via a Praat multiple forced choice interface by two rhotic Scottish-English speakers, both originally from the western Central Belt. Each judge classified the same subset of instances of prepausal postvocalic /r/ (*beer, bear, far, bar, par, purr, fur, for, bore, poor (sure, pure)*), along a 5-point continuum of auditory ‘strength’ of /r/<sup>6</sup> (ranging from graded responses such as ‘no /r/’ through ‘derhotic’, ‘alveolar’, ‘retroflex’ to full rhotic vowel, ‘schwar’). The results showed a significant association between the auditory strength of /r/ and social group, and auditory strength of /r/ and gender such that middle-class speakers showed auditorily stronger /r/ than working-class speakers, and girls showed auditorily stronger /r/ than boys (these data are shown in Figure 4 above).

---

<sup>6</sup> This was expanded to a 9-point continuum in order to take into account when both raters selected categories that were side by side on the 5-point continuum, i.e. intermediate classification categories were created.

The second — articulatory-rating — analysis of the same data involved the visual classification of the dynamic tongue gestures from the ultrasound videos. Initially, a classification system for tongue-configuration types was devised. This resulted in four categories on a scale from tip-up, through front-up to front bunched and mid-bunched, which takes the differences in configuration for retroflex /r/ and bunched /r/ as effectively lying on a continuum, e.g. Delattre & Freeman (1968) and Zhou et al. (2008). Each video was watched by the second and third authors, and the dynamic configuration of the tongue during the production of each word was noted. Examples of each are shown in the waterfall UTI diagrams in Figure 11. The articulatory-rating study also showed social stratification, with bunched variants occurring mainly in middle-class speech and tip-up variants in working-class speech.

#### FIGURE 11 ABOUT HERE PLEASE

Figure 11. Waterfall diagrams of UTI splines, sampled every 30 ms throughout words ending in /ar/, showing the dynamic movement of the tongue. Time runs in the direction of the arrows. The tongue root is to the left, tongue tip to the right. Top left: tip-up: informant LM16's utterance of *par*; Top right: front-up: LF2's utterance of *far*; Bottom left: front-bunched: EF6's utterance of *far*; Bottom right; mid-bunched: EM5's utterance of *bar*.

There was a significant correlation ( $r = 0.637$ ;  $p < 0.001$ ) between the auditory and articulatory ratings. This shows that auditorily weakened /r/, and derhoticisation in the corpus resulted from /r/ articulated with a tongue-tip raised gesture, as discussed above (§6.1), and consistent with the observation in Glasgow of working-class speakers using more taps, and being more derhotic. It also showed that the auditory continuum of rhotic-derhotic has its basis in articulation, since at the other end of the socio-articulatory continuum, the auditorily strongest postvocalic /r/ in these Eastern Central Belt speakers is the result of tongue bunching. Thus both gestural timing and tongue configuration together contribute to the percept of auditorily strong and weak rhotics in the Scottish Central belt.

#### 6.3. Accessing derhoticisation? – back to the listener

Our articulatory investigations immediately made us wonder how speakers might access, store – and reproduce – such gestures, particularly partially covert tongue-tip gestures when voicing has ceased (§6.1) or the difference between tongue tip-raising and tongue bunching (§6.2). We summarize the results of two relevant small-scale studies below.

Ashton (2011) gauged listener perceptions of articulatorily derhoticised and bunched variants of postvocalic /r/ by investigating whether they were associated with a particular geographical location or socioeconomic status. Auditory stimuli containing postvocalic /r/ were collected from the pre-existing socially-stratified UTI corpora described above and classified according to articulatory gesture (bunched approximant, apical approximant, apical derhoticised /r/ or rless – with no tongue gesture for /r/) and, in the case of derhoticised variants, strength of rhoticity. 16 participants from the Central Belt completed a computer-based subjective reaction test with randomized stimuli. Judgments were made regarding the geographical and social background of the speaker who produced each token. Bunched postvocalic /r/ was found to be strongly associated with middle-class Edinburgh speech, whereas apical approximant /r/ was associated with working-class Scottish speech, but not one particular geographical location. Derhoticised and rless realisations of postvocalic /r/ were found to be associated with Glasgow, and derhoticisation was strongly associated with working-class speech.

Lawson et al. (2011b) presents preliminary evidence for configurational lingual adaptation in Scottish postvocalic /r/ during mimicry. A male speaker, originally from the west of Scotland, was asked to mimic a number of audio stimuli extracted from the ECB08 and WL07 corpora. The articulatory gestures underlying the audio stimuli were known. His mimicked articulations were then compared to his baseline UTI recordings (only a small number of items could be compared). The mimicked data showed little adaptation of tongue configuration, but some shift in the timing of the gesture (with respect to offset of voicing) particularly when responding to the tongue-bunched auditory stimuli. It was also interesting to note that the covert, delayed apical /r/ gesture was not reproduced when mimicking the audio signal from the derhoticised utterance of *hurt*; instead the speaker produced an rless word *hut* (see Figure 12).

## FIGURE 12 ABOUT HERE PLEASE

Figure 12. Waterfall diagrams of UTI splines from the mimicking study (Lawson et al 2011b). Left: the original production of *hurt* by the mimicker, which sounds weakly rhotic. Middle: the production of the stimulus for mimicking, auditorily derhoticised *hurt*, but with covert delayed tongue-tip raising. Right: the mimicked production of *hurt*, without any tongue-tip raising, and sounding like *hut*. (Note that /t/ in these word is realized as a glottal stop.)

With respect to derhoticisation, the results confirmed that delay in the tongue tip gesture can lead to an ambiguous auditory percept not only for an analyst, but also for a derhoticising member of the Scottish speech community. Our suspicions that the acoustic signal could be difficult to parse seem plausible, though this needs more investigation, which is now underway in a systematic socio-articulatory phonetic study using mimicking in conjunction with UTI recordings. This study will allow us to develop a clearer picture of how articulatory variation spreads from speaker to hearer.

### ***7. Discussion and reflection: the sociophonology of Scottish derhoticisation***

The studies presented, both by previous scholars and ourselves, show that rhoticity in Scottish English has been eroding gradually over the 20<sup>th</sup> century for working-class speakers, and possibly for longer. This is counterbalanced by an increasingly auditorily strong rhoticity in middle-class speakers (see Lennon 2011). The changes are largely driven by sociolinguistic dynamics within this Scottish community, though there is evidence for reinforcement from an unlikely source, indirect contact with London English on TV. Describing and accounting for these changes phonetically has also been a focus— and of course — is far from complete. A number of issues arise, but we focus here on two which relate to representation, the first concentrating on analysts and how we are to deal with such labile data, the second, making suggestions about the possible mental representations held by speakers and hearers participating in these changes.

#### *7.1. Analytical representation of sociophonetic variation: the speaker-hearer triangle*

We illustrate some of the implications of our articulatory investigation for the analytical representation of variation by focusing on the middle-class, rhotic, end of continuum.

The audio-/articulatory-rating study shows clearly that auditory judgments result in auditory objects, and not the quasi-articulatory objects suggested by IPA representations. Recall that auditorily-strong approximants in middle-class speakers were consistently phonetically transcribed as ‘retroflex’ using the IPA symbol [ɭ] (e.g. Johnston 1997). However, the UTI data show that the actual configuration for these variants is likely to involve tongue-bunching, with no tip raising at all.

It is also clear that – at some level at least – the differences between tongue-tip raising and tongue bunching can be discerned by members of this speech community, since they show systematic patterning with social membership of particular subgroups. This shows that the fine-grained differences in /r/ production can be exploited and used to construct and reflect social meaning (Eckert 2008). Being an urban middle-class girl involves the use of a specific kind of auditorily-strong, bunched /r/; at the opposite pole, working-class girls in the western Central Belt are continuing to use non-rhotic and derhoticised variants. It is clear that the phonological category of /r/ in this position is closely linked with locally-situated social categories.

Moreover, these results for Scottish English are in contrast to those found for American English /r/ by Twist et al. (2007), where listeners were found to be ‘at best weakly aware’ of articulatory variation (retroflexion and bunching) in /r/ (Twist et al. 2007: 215). However, there are good reasons to assume that bunched and retroflex /r/ could be perceptually distinguished. Johnson (2011) points out that Zhou et al. (2008) identify clear acoustic differences in the frequency and trajectory of F3 and F4 between the two variants. He demonstrates their perceptual salience by showing that acoustic stimuli created with these differences can lead to differential perceptual compensation. In addition, even if acoustic equivalence is assumed for different articulatory strategies, the coarticulatory effect of these very different /r/ articulations may provide the listener with information regarding differences in underlying articulation, see Lawson et al. (2013). This suggests that rather fine-grained phonetic differences (the higher formants may often

be only weakly resonated), which are potentially accessible, have been exploited for social meaning in Scottish English, but seem to remain unattached in American English.

More generally, using these different phonetic representations of postvocalic /r/ provides a good illustration of what we call here the ‘speaker-hearer triangle’, composed of auditory, acoustic, and articulatory representations (Ogden 2009, Heselwood & Plug 2011 looks at auditory, acoustic and psychoacoustic views of rhoticity). Figure 13 shows auditorily strong postvocalic /r/: each representation gives a different picture of the ‘same’ phenomenon. In many ways each is as valid as the other, and of course, as we have seen, they are all interconnected but not necessarily in straightforward ways.

### FIGURE 13 ABOUT HERE PLEASE

Figure 13. An illustration of the ‘speaker-hearer triangle’ of auditory, acoustic, and articulatory representations of the auditorily-strong postvocalic /r/ in a middle-class Edinburgh girl’s production of the word *far*.

Ideally, representing sociophonetic variation would be able to refer to all three dimensions of the speaker-hearer triangle. Adding articulatory data can prove very fruitful (also Wright & Kerswill’s 1989 conclusions for using electropalatographic data in conjunction with auditory transcription). This can also help us to reflect on the different kinds of representation – and their intersections – that might be involved in the transmission and propagation of sociophonetic variation, which is also of crucial importance in modeling language variation and change (e.g. Marotta this volume). The traditional notion of the speaker-hearer chain (e.g. Denes & Pinson 1993) assumes that articulatory gestures from the speaker give rise to acoustic objects, which in turn become auditory objects for the listener to decode (see also Ohala e.g. 1989). How variation which appears to be so auditorily subtle, yet can be acquired and transmitted such that it can carry social meaning for a community, requires substantial further investigation.

#### *7.2. Mental representation of sociophonetic variation: a symbolic relationship?*

It is clear that the rhotic-derhotic continuum in Scottish English in the Central Belt is undergoing shifts in fine phonetic realization. It is also clear that it is impossible to

describe the scope of phonological rhoticity without reference to social factors, both macro and micro. For these speakers /r/ in this position is not just an /r/, it is always a certain kind of socially-embedded /r/; at the descriptive level it is extremely difficult to separate the phonological from the social. It is also difficult to assume that these entities do not relate to each other very closely for speaker-hearers. Data of this kind demand phonological representations which recognize the interconnected relationships between social and phonological variation which speakers in these communities need to store, control, access, and acquire.

The approach which recognizes such connections and which ‘embeds indexicality centrally within phonological knowledge’ (Foulkes & Docherty 2006: 426), is the range of theories of phonological representation grouped under the term ‘Exemplar Theory’ (e.g. Goldinger 1998, Johnson 2006, Hawkins 2003). These models share the assumption that phonological representations are based in some way on stored experiences of speech (‘exemplars’), memory clouds across which abstractions are probabilistically derived. Increasing emphasis is placed on the need for abstractions accrued from exemplar memory (corresponding to phonological categories in other perception-production models) being stored concurrently and with connections to exemplars, so-called ‘hybrid’ models (Goldinger 2007, Pierrehumbert 2006).

The results from the rhotic-derhotic continuum in Scottish English also have implications for hybrid models, particularly with respect to the relationships between phonological and social detail and abstraction. Schematic accounts of exemplar-based representations such as that by Johnson (2006) distinguish the exemplar map from accruing abstractions, but interestingly also make a separation between phonological and social categories at the abstract level. This implies that the connections between these two kinds of abstractions (as well as with others) are always made through exemplar memory. But it is clear that phonological abstractions such as ‘postvocalic /r/’, which are accessible to speakers especially through stereotypes, also relate to social abstractions at the same time. Moreover if we consider the acquisition of speech variation which is necessarily socially-embedded (e.g. Foulkes et al. 2005; Labov this volume), it seems difficult to assume that the emerging abstractions are not linked – or linkable – if only because the shared/simultaneous activation of phonological and social categories would

be so frequent. Rather these sociophonetic data, and those from many other sociophonetic studies (e.g. Foulkes et al. 2010), suggest that these abstract levels likely relate to each other directly, as the result of persistent coupling in the system.

If we make this assumption (and it seems inevitable that we must),<sup>7</sup> an analogy from ancient Greek society may be useful for considering the possible nature of the relationship between these two abstract levels. Greek *symbola* were originally two halves of the same object, each a *symbolon*, which could be fitted together for purposes of personal recognition (Herman 1987). Only later in the classical period did the meaning of the word *symbolon* shift from denoting part of a two-part tally, to tokens which could be used like tickets in exchange for goods, continued in English ‘symbolic’. The original *symbolon/symbola* relationship had two key aspects: (1) each *symbolon* could and did exist separately, for example, members of a dispersed family could keep them for a long time; but each *symbolon* was only meaningful when reunited with its partner (*symbola*). (2) *Symbola* could be formally similar, but each half could also be different from each other (Harris 2000: 23).

The relationship between phonological and social abstractions emerging from exemplar memory could be likened to the *symbolon/symbola* relationship.<sup>8</sup> Both kinds of categorization, at whatever level, can and do exist separately, both for analysts, and for speakers under particular conditions. For example, it is clearly possible to undertake separate analyses of phonological structure, or of social categorization, without reference to each other (Labov 2006). Speakers too can access phonological categories without reference to social categories, e.g. in psycholinguistic manipulation tasks, and social stereotypes can be retrieved without automatically referring to speech. But we suggest that the usual situation for speakers in daily interaction is that the social and phonological systems function in a *symbola* relationship, namely they are linked or continually linking such that “each is significant ... as a counterpart of the other” (Harris 2000: 23). Such an

---

<sup>7</sup> Keith Johnson (p.c. 2011) notes the difficulty of two-dimensional graphical representations of phenomena and processes that are a) multidimensional and b) thoroughly inter-related.

<sup>8</sup> The *symbola* relationship could also be used metaphorically to refer to the special relationships between entities; Aristotle’s account of speech and writing is given in these terms in the introduction to his *De Interpretatione*, 16<sup>a</sup>3-8.



analogy allows us to think about the social and phonological systems as having a separate, yet co-emergent relationship at the abstract level. The links themselves would be established through co-ordinated simultaneous activation, leading to persistent coupling within and across the exemplar map, and hence the entrenchment of linked/linkable social and phonological categories (this kind of modeling assumes activation and resonance discussed by Johnson 2006). At the same time, prior knowledge encapsulated in such social–phonological linkages will serve to mediate the treatment of subsequent input exemplars (Goldinger 2007).

## ***8. Conclusions***

This paper has taken an aspect of Scottish English phonology, postvocalic /r/, which appears to be changing. A strictly phonetic account of this phenomenon is not possible; social information is required even to be able to specify the kinds of phonetic variation which are emerging. Phonetic analysis carried out on socially-stratified speech data from the Scottish Central Belt shows how speakers at different ends of the social spectrum are exploiting very fine phonetic differences in coda /r/ for social ends. Conversely, such socially-informed phonetic analyses sheds light on the mechanisms of the weakening and derhoticisation of /r/, and its auditory strengthening. We discuss representations of speech from three points of the possible ‘speaker-hearer triangle’: all three give partial impressions of the phenomena. All three are needed together in order to gain an improved understanding of their nature.

Variation and change in coda /r/ is also informative for sociolinguistic theory. Unlike middle-class non-rhoticity, working-classerhoticisation in Scottish English has never been interpreted as a contact-induced change through interacting with non-rhotic English English speakers. But our results show that strong psychological engagement with a London-based television show is linked to increased r-lessness. This strongly suggests that current models of media influence, which assume that the prior knowledge of the viewer is essential, should also be extended to language, and specifically that prior sociolinguistic knowledge of the viewer may act as a sociolinguistic filter on incoming

media language – leading to decay or enhancement depending on the degree of social relevance and linguistic congruence with the speaker/viewer's system.

Overall, we have learnt a lot, but there is still much more to discover, both about this particular phenomenon, and about some of the wider issues which it exemplifies, for example:

- What has happened in real time over the past century? Are we witnessing language change, and if so how fast or gradual is this? Only empirical study of real-time data can begin to answer this question.

- How can we objectively describe and assess derhoticisation? We need improved understanding of the acoustics, and the psychoacoustics, of rhoticity.

- How do changes in coda position relate to those in onset position? Our study focuses on coda /r/, particularly in utterance-final position. We have noted that this location seems to be particularly salient socially. More work needs to be carried out – like that of Pukli and Jauriberry (2011) – which analyses /r/ in all positions.

- How does subtle articulatory variation of this kind get transmitted? Modeling mechanisms of language variation and change rely on a much improved understanding of the relationships between speakers and hearers, and in particular, how hearers may respond to input from speakers at the level of articulation.

- How do speakers phonetically and socially decode speech experienced without the possibility of interaction? This area is virtually unresearched, but needs to be explored empirically if we are to make progress in understanding how engaging with the broadcast media relates to spoken language in the community.

Our current and future research, with each other and other colleagues, aims to try to tackle some of these questions. But it is now clear to us, after working on this phenomenon for over 15 years, that what appears to be the answer is usually the starting point for more questions: in fact this particular sociophonetic journey has only just begun.

Aitken, Adam J. 1984. "Scots and English in Scotland". *Language in the British Isles* ed. by Peter Trudgill, 517-532. Cambridge: Cambridge University Press.

- Anthony, A., D. Bogle, T.T.S. Ingram & M.W. McIsaac. 1971. *The Edinburgh Articulation Test*. Edinburgh: Churchill Livingstone.
- Articulate Instruments Ltd. (2011). *Articulate Assistant Advanced User Guide: Version 2.13*. Edinburgh: Articulate Instruments Ltd.
- Ashton, Lindsay. 2011. *Perception of /r/ in Scottish English*. Undergraduate Dissertation, Edinburgh, Queen Margaret University.
- Bucholtz, Mary. 2009. "From stance to style: gender, interaction, and indexicality in Mexican immigrant youth slang". *Stance. Sociolinguistic Perspectives on Stance* ed. by Alexandra Jaffe, 146-170. New York: Oxford University Press.
- Carey, Emma. 2010. *Cross Accent Perception of Standard Southern British English and Glasgow English*, Undergraduate Dissertation, University of Glasgow.
- Coupland, Nikolas. 2007. *Style: Language variation and identity*. Cambridge: Cambridge University Press.
- Denes, Peter B. & Elliot N. Pinson, E. 1993. *The Speech Chain: The Physics and Biology of Spoken Language*. Oxford: W.H. Freeman and Company.
- Delattre, Pierre & Donald C. Freeman. 1968. "A dialect study of American r's by x-ray motion picture". *Linguistics* 44. 29-68.
- Dobson, Eric J. 1957. *English Pronunciation 1500-1700*. Oxford: Clarendon Press.
- Docherty, Gerard J. & Paul Foulkes. 1999. "Derby and Newcastle: instrumental phonetics and variationist studies". *Urban voices: variation and change in British accents* ed. by Paul Foulkes & Gerard J. Docherty, 47-71. London: Arnold.
- Douglas, Fiona M. 2009. *Scottish Newspapers, Language and Identity*. Edinburgh: Edinburgh University Press.
- Durand, Jacques. 2004. "English in early 21st century Scotland: a phonological perspective". *La Tribune Internationale des Langues Vivantes* 36. 87-108.
- Eckert, Penelope. 2008. "Variation and the indexical field". *Journal of Sociolinguistics* 12(4). 453-76.
- Foulkes, Paul and Gerard J. Docherty. 2006. "The social life of phonetics and phonology". *Journal of Phonetics* 34. 409-38.
- Foulkes, Paul, Gerard J. Docherty & Dominic J.L. Watt. 2005. "Phonological variation in child directed speech". *Language* 81. 177-206.

- Foulkes, Paul, James M. Scobbie & Dominic J.L. Watt. 2010. "Sociophonetics". *Handbook of Phonetic Sciences* (2nd ed.) ed. by William Hardcastle, John Laver & Fiona Gibbon, 703-754. Oxford: Blackwell.
- Goldinger, Stephen D. 1998. "Echoes of echoes? An episodic theory of lexical access". *Psychological Review* 105. 251-79.
- Goldinger, Stephen D. 2007. "A complementary-systems approach to abstract and episodic speech perception". *Proceedings of the 16th International Congress of Phonetic Sciences* ed. by Jürgen Trouvain & William J. Barry, 49-54 (ID 1781). <http://www.icphs2007.de/>
- Grant, William. 1914. *The Pronunciation of English in Scotland*. Cambridge: Cambridge University Press.
- Grant, William and James M. Dixon. 1921. *A Manual of Modern Scots*. Cambridge: Cambridge University Press.
- Gunter, Barrie. 2000. *Media research methods: Measuring audiences, reactions and impact*. London: Sage.
- Hall, Stuart. 1980. "Encoding/decoding". *Culture, Media, Language: Working Papers on Cultural Studies*, 1972-79, 128-38. London: Hutchinson.
- Harris, Roy. 2000. *Rethinking writing*. London: Athlone Press.
- Hawkins, Sarah. 2003. "Roles and representations of systematic fine detail in speech understanding". *Journal of Phonetics* 31. 373-405.
- Herman, Gabriel. 1987. *Ritualised Friendship and the Greek City*. Cambridge: Cambridge University Press.
- Heselwood, Barry and Leendert Plug. 2011. "The role of F2 and F3 in the perception of rhoticity: Evidence from listening experiments". *Proceedings of the XVIII International Congress of Phonetic Sciences*, 867-70. [http://www.icphs2011.hk/ICPHS\\_CongressProceedings.htm](http://www.icphs2011.hk/ICPHS_CongressProceedings.htm)
- Jaffe, Alexandra. 2009. "Introduction: The sociolinguistics of stance". *Stance. Sociolinguistic Perspectives on Stance* ed. by Alexandra Jaffe. New York: Oxford University Press.
- Johnson, Keith. 2006. "Resonance in an exemplar lexicon: The emergence of social identity and phonology". *Journal of Phonetics* 34. 485-499.

- Johnson, K. (2011), 'Retroflex versus bunched /r/ in compensation for coarticulation', *UC Berkeley Phonology Lab Annual Report*. 114-27.
- Johnston, Paul. 1983. "Irregular style variation patterns in Edinburgh speech". *Scottish Language* 2: 1-19.
- Johnston, Paul . 1985. "The rise and fall of the Morningside/Kelvinside accent". *Focus on Scotland* ed. by Manfred Gorlach, 37-56. Amsterdam: Benjamin.
- Johnston, Paul. 1997. "Regional Variation". *The Edinburgh History of Scots* ed. by Charles Jones, 433-513. Edinburgh: Edinburgh University Press.
- Kuhl, Patricia. 2010. "Brain mechanisms in early language acquisition". *Neuron* 67. 713-775.
- Labov, William. 1972. *Sociolinguistic patterns*. Oxford: Blackwell.
- Labov, William. 1994. *Principles of Linguistic Change, vol. I, Internal Factors*. Oxford: Blackwell.
- Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- Lawson, Eleanor, Jane Stuart-Smith & James M. Scobbie. 2008. "Articulatory insights into language variation and change: Preliminary findings from an ultrasound study of derhoticisation in Scottish English". *University of Pennsylvania Working Papers in Linguistics* 14 (2). 101-110.
- Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2011. "The social stratification of tongue shape for postvocalic /r/ in Scottish English". *Journal of Sociolinguistics* 15(2). 256-268
- Lawson, Eleanor, Jane Stuart-Smith, James M. Scobbie, Malcah Yaeger-Dror & Margaret MacLagan. 2011a. "Liquids". *Sociophonetics: A Student's Guide* ed. by Malcah Yaeger-Dror and Marianna di Paolo, 72-86. London: Routledge.
- Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2011b. "A single case study of articulatory adaptation during acoustic mimicry". *Proceedings of the XVIII International Congress of Phonetic Sciences*, 1170-1173.  
[http://www.icphs2011.hk/ICPHS\\_CongressProceedings.htm](http://www.icphs2011.hk/ICPHS_CongressProceedings.htm)

- Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2013. "Bunched /r/ promotes vowel merger to schwa: An Ultrasound Tongue Imaging study of Scottish sociophonetic variation." *Journal of Phonetics* 41. 298-210
- Lawson, Eleanor, James M. Scobbie & Jane and Stuart-Smith. 2013, forthcoming. "A socio-articulatory study of Scottish rhoticity". *Sociolinguistics in Scotland* ed. by Robert Lawson, Basingstoke: Palgrave.
- Lennon, Robert. 2011. *A real-time study of rhoticity in Glaswegian between 1997 and 2011*. Undergraduate Dissertation, University of Glasgow
- Llamas, Carmen. 2010. "Convergence and divergence across a national border". *Language and Identities* ed. by Carmen Llamas & Dominic Watt, 227-236. Edinburgh: Edinburgh University Press.
- Llamas, Carmen, Dominic Watt & Daniel Ezra Johnson. 2009. "Linguistic Accommodation and the Salience of National Identity Markers in a Border Town". *Language and Social Psychology* 28. 381-407.
- Macafee, Caroline. 1983. *Varieties of English around the World: Glasgow*. Amsterdam: John Benjamins.
- MacAllister, Anne H. 1938. *A Year's Course in Speech Training*. London: Hodder and Stoughton.
- Macaulay, Ronald K.S. 1977. *Language, Social Class and Education: A Glasgow Study*. Edinburgh: Edinburgh University Press.
- MacFarlane, Andrew E. & Jane Stuart-Smith. 2012. "'One of them sounds sort of Glasgow Uni-ish': Social judgements and fine phonetic variation in Glasgow". *Lingua* 122(7). 764-778.
- Ohala, John J. 1989. "Sound change is drawn from a pool of synchronic variation". L. E. Breivik & E. H. Jahr (eds.), *Language Change: Contributions to the study of its causes* ed. by Leiv E. Breivik & Ernst H. Jahr, 173-198. Berlin: Mouton de Gruyter.
- Pierrehumbert, Janet. 2006. "The next toolkit". *Journal of Phonetics* 34. 516-530.
- Plug, Leendert & Richard Ogden. 2003. "A parametric approach to the phonetics of postvocalic /r/ in Dutch". *Phonetica* 60. 159-86.

- Pukli, Monika and Thomas Jauriberry. 2011. "Language change in action – variation in Scottish English". *Recherches Anglaises et Nord-Américaines* 44. 83-100.
- Richmond, Suzanne. 2013. *A Real-Time Sociophonetic Study of Scottish English: The Realisation of /r/ across a Century*. Unpublished Honours Dissertation, University of Glasgow.
- Romaine, Suzanne. 1978. "Postvocalic /r/ in Scottish English: sound change in progress?". *Sociolinguistic patterns in British English* ed. by Peter Trudgill, 144-57. London: Edward Arnold.
- Scobbie, James M. 2006. "R as a variable". *Encyclopedia of Language and Linguistics* ed. by Keith Brown, 334-344. Amsterdam: Elsevier.
- Scobbie, James M. 2007. "Biological and social grounding of phonology: Variation as a research tool". *Proceedings of the 16th International Congress of Phonetic Sciences* ed. by Jürgen Trouvain & William J. Barry, 211-214 (ID 1765). <http://www.icphs2007.de/>
- Scobbie, James M., Olga B. Gordeeva & Benjamin Matthews. 2007. "Scottish English Speech Acquisition". *The International Guide to Speech Acquisition* ed. by Sharynne McLeod, 221-240. Thomson Delmar Publishing.
- Scobbie, James M. and Koen Sebregts. 2011. "Acoustic, articulatory and phonological perspectives on rhoticity and /r/ in Dutch (Eds.) *Interfaces in Linguistics: New Research Perspectives* ed. by Raffaella Folli & Christiane Ulbrich, 257-277. Oxford: Oxford University Press.
- Scobbie, James M. & Jane Stuart-Smith. 2008. "Quasi-phonemic contrast and the fuzzy inventory: examples from Scottish English". *Contrast: Perception and Acquisition. Selected papers from the Second International Conference on Contrast in Phonology* ed. by Peter Avery, Elan B. Dresher & Keren Rice, 87-113. Berlin: Mouton de Gruyter.
- Scobbie, James M. & Jane Stuart-Smith. 2012. "Socially-stratified sampling in laboratory-based phonological experimentation". *The Oxford Handbook of Laboratory Phonology* ed. by Abigail C. Cohn, Cécile Fougeron & Marie Huffman, 608-621. Oxford: Oxford University Press.

- Speitel, Hans-Henning & Paul Johnston. 1983. "A Sociolinguistic investigation of Edinburgh speech". Unpublished final report (Grant No. 000230023) for the Social Science Research Council (Great Britain).
- Sproat, Richard & Osamu Fujimura. 1993. "Allophonic variation in English /l/ and its implications for phonetic implementation". *Journal of Phonetics* 21. 291-311.
- Stevens, Kenneth. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press
- Stuart-Smith, Jane. 1999. "Glasgow: accent and voice quality". *Urban voices: variation and change in British accents* ed. by Paul Foulkes & Gerard Docherty, 201-222. London: Arnold.
- Stuart-Smith, Jane. 2003. "The phonology of Modern Urban Scots". *The Edinburgh Companion to Scots* ed. by John Corbett, Derrick McClure & Jane Stuart-Smith, 110-137. Edinburgh: Edinburgh University Press.
- Stuart-Smith, Jane. 2006. "The influence of media on language". *The Routledge Companion to Sociolinguistics* ed. by Carmen Llamas, Louise Mullany, Peter Stockwell, 140-148. London: Routledge.
- Stuart-Smith, Jane. 2007. "A sociophonetic investigation of postvocalic /r/ in Glaswegian adolescents". *Proceedings of the 16th International Congress of Phonetic Sciences* ed. by Jürgen Trouvain & William J. Barry, 211-214 (ID 1307). <http://www.icphs2007.de/>
- Stuart-Smith, Jane & Claire Timmins. 2010. "The role of the individual in language change". *Language and Identity* ed. by Carmen Llamas & Dominic Watt, 39-54. Edinburgh: Edinburgh University Press.
- Stuart-Smith, Jane, Rachel Smith, Tamara Rathcke, Francesco Li Santi & Sophie Holmes. 2011. "Responding to accents after experiencing interactive or mediated speech". *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, 1914-1917. <http://www.icphs2011.hk>
- Stuart-Smith, Jane, Claire Timmins & Fiona Tweedie. 2007. "Talkin' Jockney: Accent change in Glaswegian". *Journal of Sociolinguistics* 11. 221-61.
- Stuart-Smith, Jane, Claire Timmins, Gwilym Pryce & Barrie Gunter. 2013. "Television is also a factor in language change: Evidence from an urban dialect". *Language*. 89 (3). 1-36.
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.



- Twist, Alina, Adam Baker, Jeff Mielke & Diana Archangeli. 2007. "Are 'covert' /r/ allophones really indistinguishable?". *University of Pennsylvania Working Papers in Linguistics* 13(2). 207-216.  
<http://repository.upenn.edu/pwpl/vol13/iss2/16>
- Wells, John C. 1982. *Accents of English*. Cambridge, Cambridge University Press.
- Whitaker, Charles W.A. 1996. *Aristotle's De Interpretatione: Contradiction and Dialectic*. Oxford: Clarendon.
- Williams, Irene F. 1909. *Phonetics for Scottish students: the sounds of polite Scottish described and compared with those of polite English*. Glasgow: J. MacLehose.
- Williams, Ann & Paul Kerswill. 1999. "Dialect levelling: change and continuity in Milton Keynes, Reading and Hull". *Urban voices: variation and change in British accents* ed. by Paul Foulkes & Gerard J. Docherty, 141-162. London: Arnold.
- Wright, Susan & Paul Kerswill. 1989. "Electropalatography in the study of connected speech processes". *Clinical Linguistics and Phonetics* 3. 49-57.
- Yaeger-Dror, Malcah, Tyler Kendall, Paul Foulkes, Dominic Watt, Jillian Oddie, Phil Harrison & Colleen Kavenagh. 2009) "Perception of r-fulness by trained listeners". Paper presented at NWAV 37, Houston, TX.
- Zhou, Xinhui, Carol Espy-Wilson, Mark Tiede, Suzanne Boyce, Christy Holland & Ann Choe. 2008. "A magnetic resonance imaging-based articulatory and acoustic study of 'retroflex' and 'bunched' American English /r/". *Journal of the Acoustical Society of America* 123(6). 4466-4481.

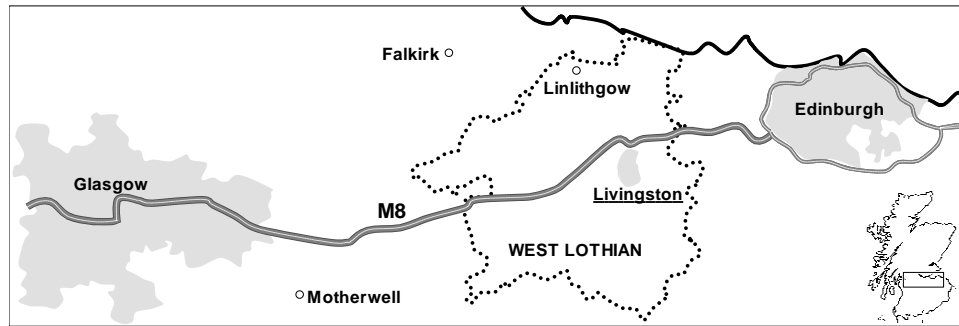


Figure 1: The Central Belt of Scotland (see inset) showing the cities of Glasgow on the west, Edinburgh on the East, and Livingston in between (from Lawson et al 2008).

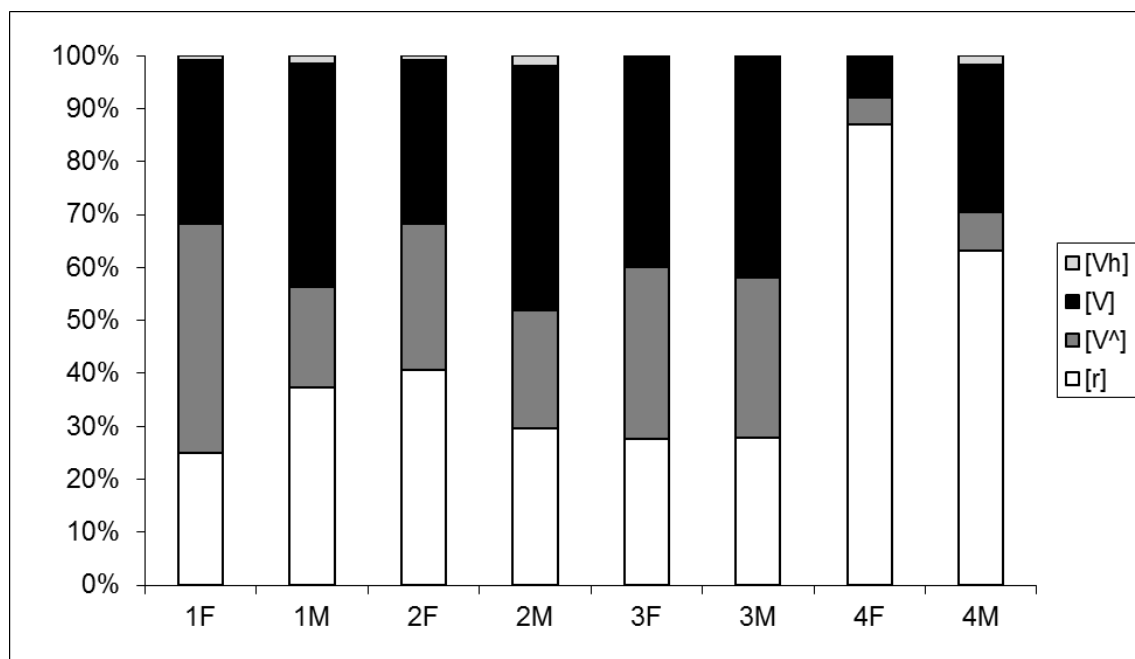


Figure 2. Distribution of variants of postvocalic /r/ in 48 speakers of Glaswegian in 2003, n = 1889. M = male, F = female; 1 = 10-11 years; 2 = 12-13 years; 3 = 14-15 years; 4 = 40-60 years. [r] = articulated variants of /r/; [V^] = vowels with audible pharyngealisation/uvularisation; [V] = plain vowel; [Vh] = vowel followed by audible frication.

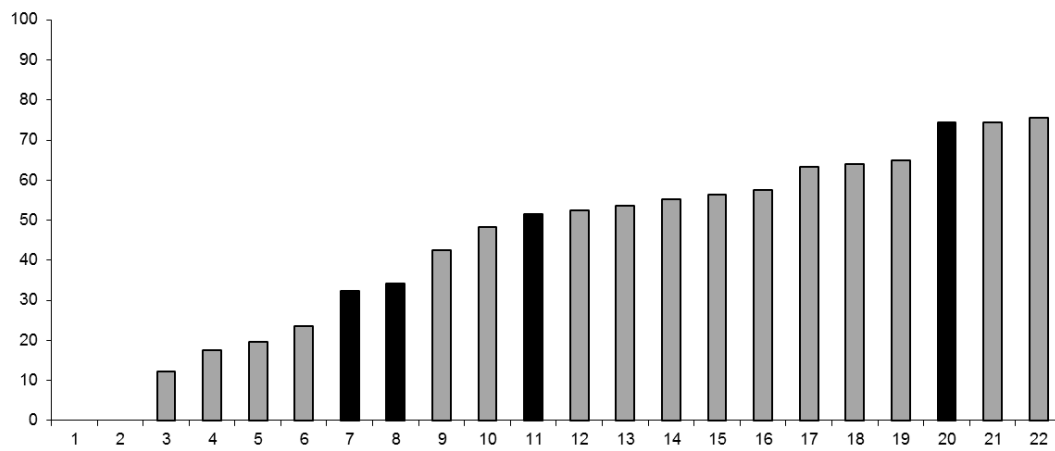
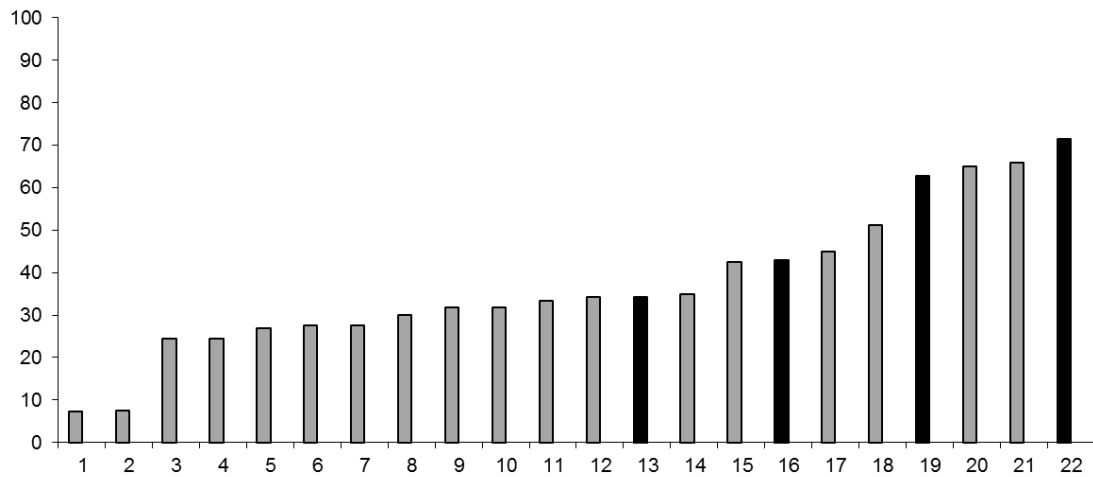


Figure 3: Percentage of the plain vowel variant for coda /r/ used by 42 speakers, 36 recorded in 2003 (pale bars) and 8 recorded in 1997 (dark bars). The top chart shows female speakers, the bottom, male speakers.

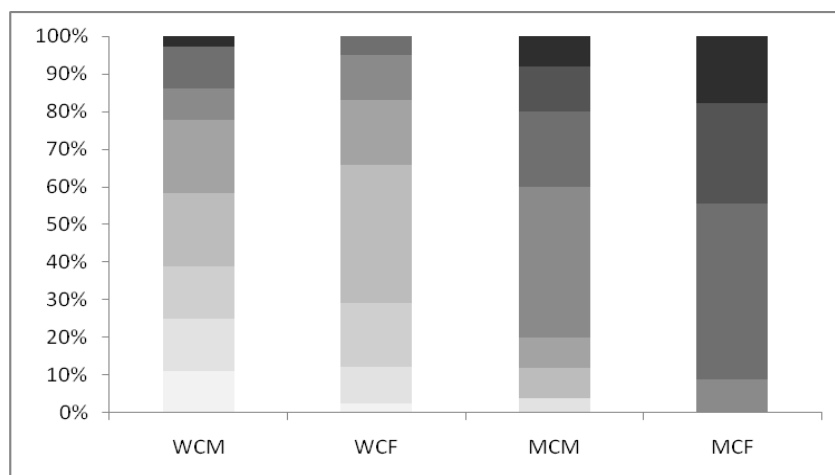


Figure 4: Bar graph showing the percentage of auditory variants used by each socioeconomic and gender group in the ECB08 corpus. WC/MC = working/middle-class; M/F = male/female. Paler grey segments represent rless and weakly rhotic variants, while darker grey segments represent strongly rhotic variants. N=139. From Lawson et al (2011), Fig. 2.

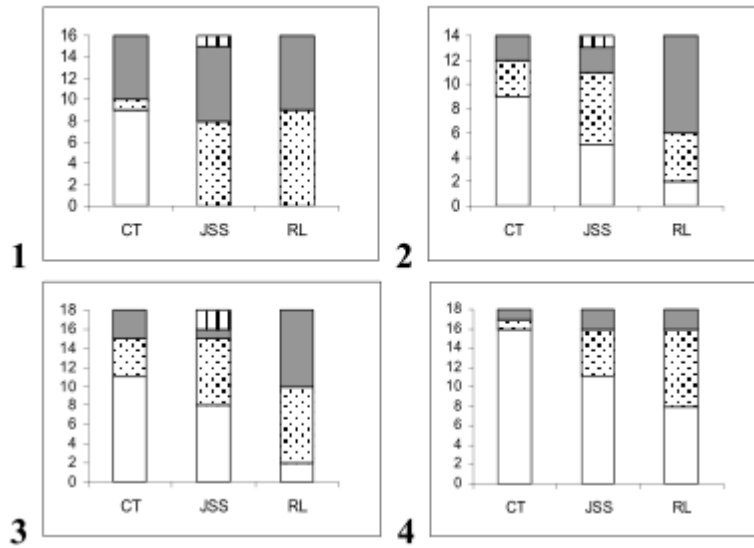
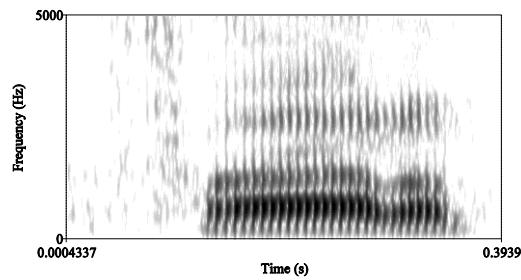
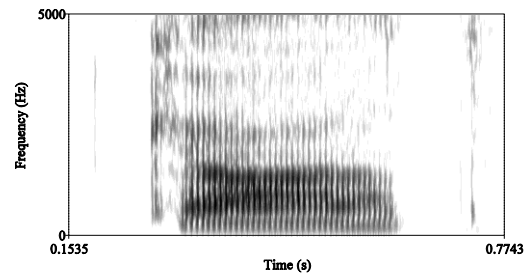


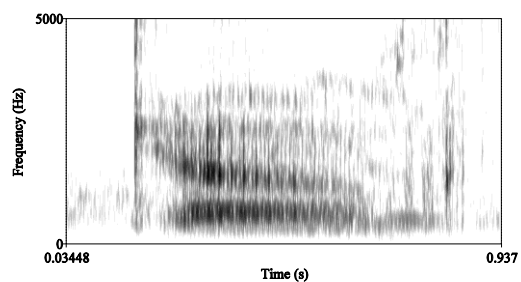
Figure 5: Results of the auditory transcription of postvocalic /r/ in word-list data read by 12 male Glaswegian working-class speakers, organised into four age groups (1= 10-11, 2= 12-13, 3=14-15, 4= 40-60). The judgments of the three transcribers (CT, JSS, RL) are shown in each chart from left to right. White = articulated /r/, spotted = pharyngealised/uvularised vowels, grey = plain vowels, striped = vowels followed by [h] or [ɦ], from Stuart-Smith (2007: 1308).



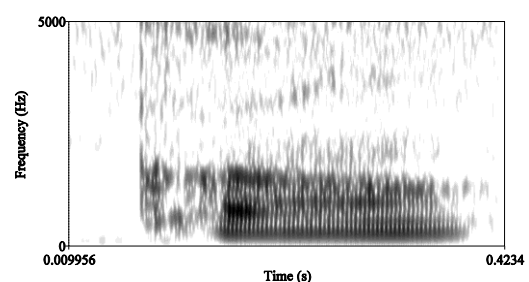
(a) *farm* with tap (adult male)



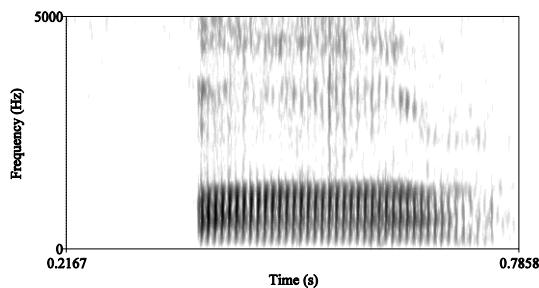
(d) *card* with pharyngealized/uvularized vowel (12 yr-old boy)



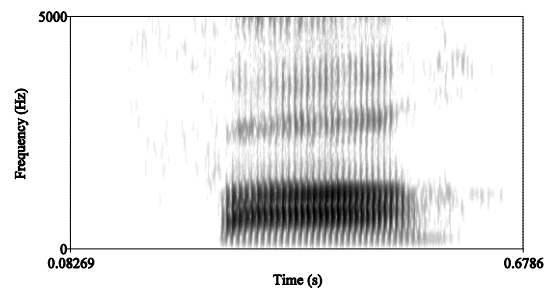
(b) *car* with trill (adult male)



(e) *car* with plain vowel (11 yr-old boy)



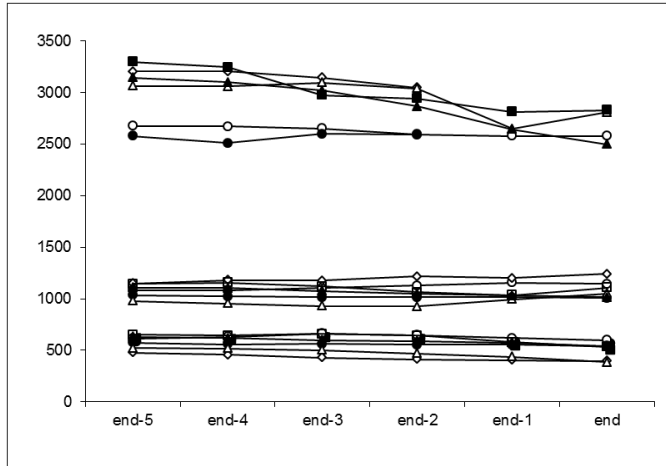
(c) *far* with weak approximant (14 yr-old boy)



(f) *far* with vowel followed by weak frication (14 yr-old boy)

Figure 6: Spectrograms illustrating the four auditory variant categories shown in Figure 5. Articulated /r/ is shown on the left in (a)-(c); vowel variants are on the right – pharyngealized/uvularized vowel (d), plain vowel (e), and vowel followed by weak frication (f). All recordings were made in 2003 in Glasgow.

(a)



(b)

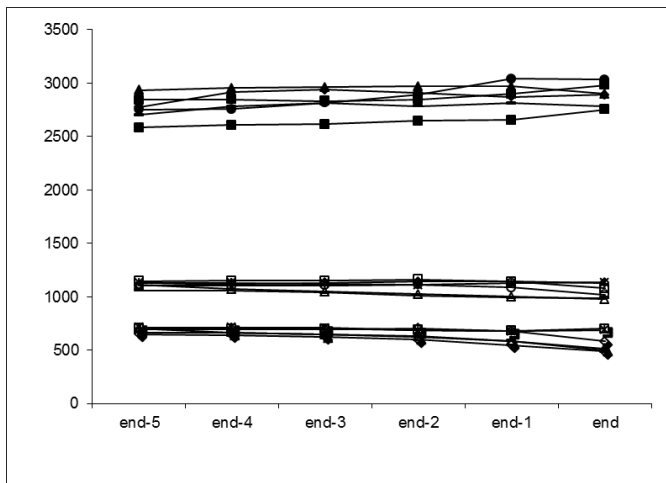


Figure 7: Handcorrected time-normalized formant tracks taken at the end of the vocalic portion and for each of the five preceding pulses, for the first three formants for two speakers: a) 14 year-old boy heard as rhotic, shows slight dip in a high F3 in most words with /r/ (this boy produced *far*, Fig 6c). b) 14 year-old boy heard with mainly pharyngealized vowels for words with /r/, shows high, flat or rising F3, with weak amplitude (this boy produced *far*, Fig 6f).



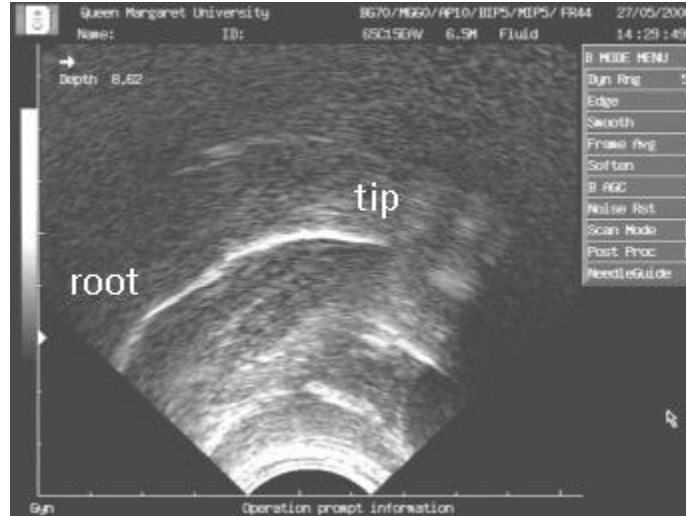


Figure 8: Midsagittal image of the tongue surface produced using a Concept M6 medical ultrasound machine. The tongue root is to the left of the image and the tip is to the right of the image.

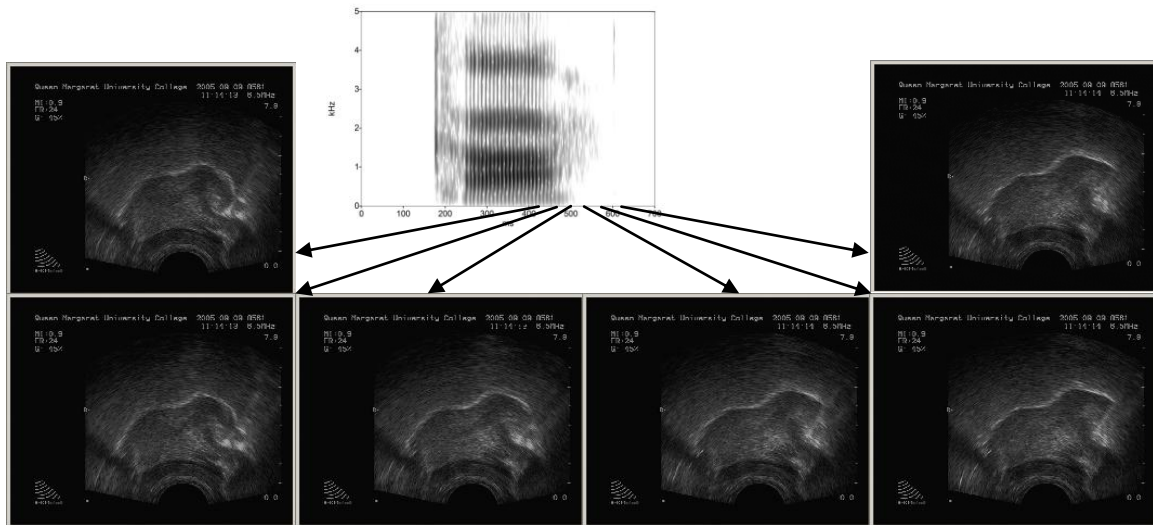


Figure 9: Key UTI frames of an adult male speaker from West Lothian, saying *car* showing a covert tip-raising gesture in the production of coda /r/. The ultrasound images correspond to the time point of the spectrogram. Moving through the frames, it is clear that the tongue front and tip begins to rise after voicing has ceased, and achieves its maximum raising well after.

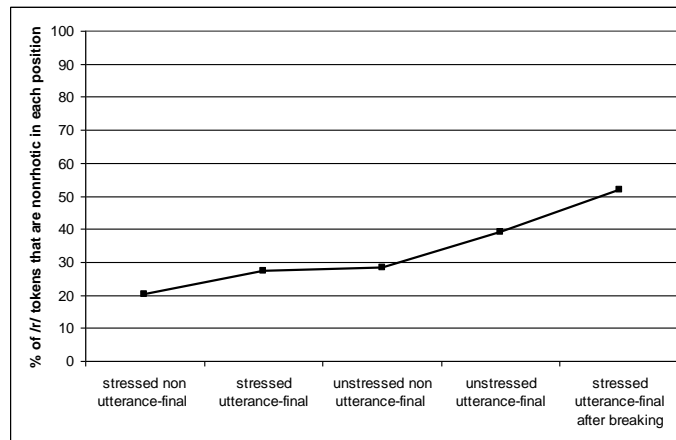


Figure 10: Percentage of (un)stressed tokens in utterance-final and non utterance-final position that were audibly nonrhymic. n=1248. From Lawson et al (2008).

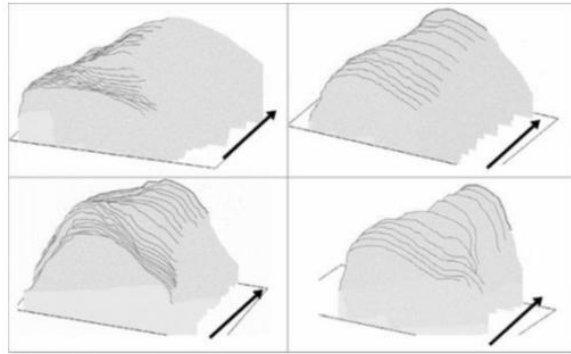


Figure 11: Waterfall diagrams of UTI splines, sampled every 30 ms throughout words ending in /ar/, showing the dynamic movement of the tongue. Time runs in the direction of the arrows. The tongue root is to the left, tongue tip to the right. Top left: tip-up: informant LM16's utterance of *par*; Top right: front-up: LF2's utterance of *far*; Bottom left: front-bunched: EF6's utterance of *far*; Bottom right; mid-bunched: EM5's utterance of *bar*.



Figure 12: Waterfall diagrams of UTI splines from the mimicking study (Lawson et al 2011b). Left: the original production of *hurt* by the mimicker, which sounds weakly rhotic; Middle: the production of the stimulus for mimicking, auditorily derhoticised *hurt*, but with covert delayed tongue-tip raising. Right: the mimicked production of *hurt*, without any tongue-tip raising, and sounding like *hut*. (Note that /t/ in these word is realized as a glottal stop.)

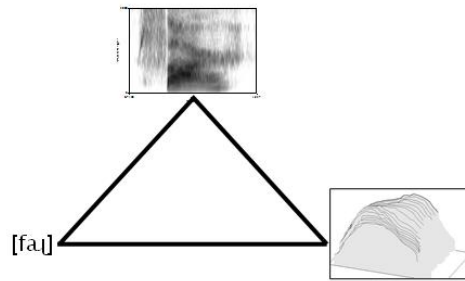


Figure 13: An illustration of the ‘speaker-hearer triangle’ of auditory, acoustic, and articulatory representations of the auditorily-strong postvocalic /r/ in a middle-class Edinburgh girl’s production of the word *far*.